**DevOps tools**

A core concept within the technical fraternity, DevOps is where the software development and operations meet to ensure continuous integration (CI) and delivery (CD). For those of you who are not familiar with the concept, DevOps, if put in layman’s term, is:

A combination of processes in which both the software engineers and operation engineers work as a unified front throughout the entire software development cycle, starting from ideation to production support.

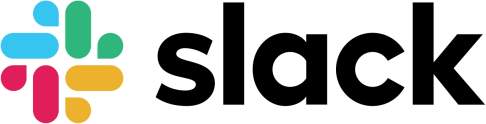
Initially, the development and the service sides were considered as two separate entities and were responsible for handling different dimensions in a software cycle. However, there were numerous siloed challenges such as increased development time, low throughput, etc. faced due to this.

Increased production time, improved communication, reduced cost, and deployment failures are the key motivators behind DevOps. For a company to completely harness the power of DevOps, they need to take care of five major processes, i.e. Communication, CI/CD, Configuration Management, Security and Monitoring & Alerting.

Top 10 DevOps Tools

As technology advances, numerous DevOps tools have been developed to make collaboration and development easier. To help you refine your DevOps strategy, we mention 10 top DevOps tools which you should use in the year 2020.

1. Slack



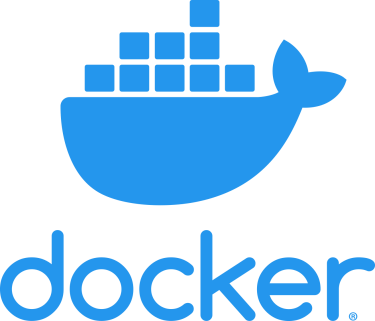
Launched in the year 2013, Slack is still one of the top communication tools used by teams for effective collaboration on projects. This tool in the DevOps arsenal is used by technical organizations across the globe to tear down barriers and offer all team members a clear insight into the workflow. One exciting feature about Slack is that it allows developers to collaborate using tool chains in the same environment they are communicating with other maintenance and service members.

2. Jenkins



An open-source continuous integration server, Jenkins automates the complete build cycle of a software project. The USP of this tool is the Pipeline feature it offers, which can be utilized by the developers to automatically commit code into the repository, run test cases, as well as fetch reports obtained after testing. This highly customizable tool provides instant feedback and hence will warn you if a particular sprint is leading to a broken build or is harming the same. Most of the tasks and tools involved in [SDLC](https://hackr.io/blog/sdlc-methodologies) can be automated using Jenkins, allowing team members to increase their throughput.

3. Dockers



Docker is a tool which is at the center of containerization, a trend which is quickly gaining momentum in the IT world. Docker allows secure packaging, deploying and running of applications irrespective of the running environment. Every application container contains the source code, supporting files, run time, system config files, etc. responsible for application execution. Using the Docker Engine, the containers can be accessed which in turn can execute applications in a remote environment. The app has allowed organizations to reduce infrastructure costs. According to a report, 2 out of 3 companies who have tried this application have adopted it within 30 days of using it.

4. Phantom



Security of software is one of the prime concerns of any DevOps team. As such, the Phantom tool comes as a great help to developers who wish to build a defensible infrastructure from the very beginning of SDLC. Using the phantom tool, you can collaborate in a centralized environment on an incident and be aware of the rising security threats at the same time. The tool further gives DevOps professionals an option to mitigate such risks instantaneously using techniques such as file detonation, device quarantine, etc.

5. Nagios



Similar to Phantom, Nagios is also a monitoring tool which tends to keep tabs on the applications, servers as well as your overall business infrastructure. The tool comes in as a great help for large organizations which have a countless number of circuitry (routers, servers, switches, etc.) in the backend. It alerts the users in case a particular fault occurs on the backend or any device fails. It also regularly maintains a performance chart and monitors trends to alert the user of a possible failure which may occur.

6. Vagrant



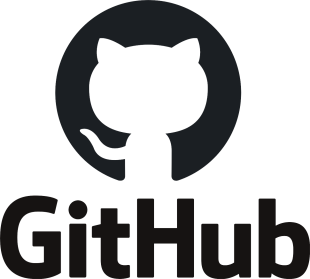
A vagrant is a tool for managing and working with virtual machines in a single workflow. Using Vagrant, team members can share software running environment and can test applications faster without wasting time on setting up configurations. The tool ensures that the environment for a particular project remains the same across every developer’s machine, and the excuse of ‘runs-on-my-system’ can be thrown out of the window.

7. Ansible



Ansible is one of the simplest yet effective IT orchestration and configuration management tools available in the market. Compared to its competitors such as Puppet and Chef, which are loaded with features, Ansible offers a softer outlook and doesn’t hog on your device’s resources in the background. This tool is primarily utilized for pushing new changes within the existing system, as well as configuring newly deployed machines. Lowering the costs of infrastructure and increasing the replication speed of scalability are just two of the reasons which have made this an absolute favourite amongst IT companies.

8. GitHub



Launched in the year 2000, GitHub remains as one of the top DevOps tools for easy collaboration. Using this tool, developers can make rapid iterations to the code, the notification of which is sent instantaneously to other team members. In case of any error or fallout, immediate rollbacks can be done to the previous version within seconds, thanks to the branched history of changes which are stored contiguously within the tool.

9. Sentry



Sentry, a tool used by companies such as Uber and Microsoft, is one of the best DevOps tools for error or bug detection. This free tool supports languages such as Ruby, IOS, JavaScript, etc. and further has inbuilt SDKs which can be customized for supporting most languages and frameworks. The tool continuously scans lines of code across the entire system and sends notifications if it finds an error or problem. Not only does it highlight the problem, but offers a choice of possible solutions which can be incorporated with a single click.

10. BitBucket

/

Similar to GitHub, BitBucket is also a tool which helps manage project code throughout the software development cycle. While GitHub still ranks as the top repository, people are shifting to BitBucket due to its reduced cost as well as the feature of the private repository (a feature only available in the paid variant of GitHub). While the core functionality of BitBucket resembles that of GitHub, features such as easy integration with Jira and Trello, inbuilt CI/CD functionality tend to give this tool by Atlassian an edge.

Conclusion

So, these were the top 10 DevOps tools which are being increasingly adopted by companies across the globe. If you think that there are any better tools which we missed, don’t forget to mention them in the comment section below.

**DataOps tools**

# Introduction

Now that organizations are dealing with big data on a day-to-day basis to generate useful insights, we require more efficient software/data development lifecycles.

The era of big data calls for some powerful data operation tools which can automate processes and reduce the cycle time of the data analytics for enormous datasets. For this purpose, the concept of DataOps is what works as the solution. It is a process-oriented methodology that monitors and controls the data analytics pipeline by using Statistical Process Controls.

In this article, we will discuss the role of some great data ops tools. So without further ado, let’s start.

# Why are DataOps Tools Important

[DataOps](https://logitanalytics.com/what-is-dataops-and-how-is-it-changing-the-data-world/) is not just about managing data pieces, it is more about delivering business value. This methodology is a combination of data related elements and softwares which together run the business operations. It is built with the use of DevOps – a widely accepted practice for accelerating software development – in a more sophisticated manner.

With DataOps Tools, you can deliver new and existing data services more quickly despite the changing semantics and infrastructures of data environments. The DataOps tools also help applications in interacting more easily while working with dynamic technologies. Furthermore, the tools transform stodgy BI into democratized and real-time analytics capability which obviously unlocks a bigger potential.

# 9 Great DataOps Tools

Now that we understand what DataOps Tools are and why they are important, let’s discuss some most popular tools:

## Data Pipeline Tools

Simply put, [data pipelines](https://www.theseattledataguy.com/airbnbs-airflow-versus-spotifys-luigi/) provide organizations access to well-structured, reliable datasets so as to extract useful analytics and insights. This helps get data from operational and application systems into data warehouses analytical systems. Some of the most popular data pipelining tools include:

### Genie

**Website Link:** <https://netflix.github.io/genie/>

Developed by Netflix, the DataOps tool is an open-source engine that offers distributed job orchestration services. This tool provides RESTful APIs for developers who wish to run a wide range of jobs with Big Data, such as Hive, Hadoop, Presto, and Spark. Genie also provides APIs for metadata management in distributed processing clusters.

### Piper

**Website Link:** <https://www.piperr.io/>

Piper is a package of Machine Learning based DataOps tools that enable organizations to read data more smoothly and efficiently. This solution exposes data through a set of APIs which integrate easily with digital assets of the organization. Furthermore, it merges batch and real-time to offer the best of data technologies along with detailed support. With a focus on AI, Pipper allows companies to minimize turnaround time of data operations and manages a complete software development lifecycle through its prepackaged data apps.

### Airflow

**Website Link:**<https://airflow.apache.org/>

Apache Airflow is an open-source DataOps platform that manages complex workflows in any organization by considering data processes as DAG (Directed Acyclic Graphs). This took was first designed by Airbnb to schedule and monitor their workflows. Now organizations can utilize this open-source tool to manage their data process on macOS, Linux, and Windows.

## Automated Testing Tools

The second category of DataOps tools covers automated testing. Simply put, automated testing tools test and compare the actual outcomes of a software technique, versus the expected outcome. These tests are applied to repetitive tasks to identify the best methods.

### Naveego

**Website Link:**<https://www.naveego.com/>

Naveego is a cloud data integration platform that allows businesses to reach accurate business decisions by integrating all company data in a regular business-centric format. This tool cleans stored data and makes it analytics-ready for data scientists. With Naveego, you can conveniently monitor and validate all your company’s stored data with security.

### FirstEigen

**Website Link:**<http://firsteigen.com/>

FirstEigen is a platform including Machine Learning tools that provide big data quality validation and data matching on the basis of self-learning. This platform learns about data quality behaviors and models using advanced ML techniques and then tests big data with just three clicks. With FirstEigen, organizations can ensure accuracy, completeness, and sanctity of their data as it moves via multiple IT platforms.

### RightData

**Website Link:**<https://www.getrightdata.com/>

RightData is a self-service group of applications designed for achieving data quality assurance, integrity audit and continuous control along with automated validation. This suite is best suited for organizations seeking tools with automated testing and reconciliation capabilities. With RightData, you can achieve testing for Data Migration, Database Upgrades, DAP, BI, reports, and much more.

## Data Science Model Deployment Tools

Model deployment is basically a method in which you integrate the AI or ML data model into any existing production environment so as to make business decisions based on the data sets. This is usually the last step in the model lifecycle and therefore, it is very crucial.

### Badook

**Website Link:** <https://badook.ai/index.html>

Badook is a popular tool among data scientists since it allows them to write automated tests for datasets used in training/testing data models. This tool not only allows them to validate data automatically but also reduces turnaround time for generating insights.

### DataKitchen

**Website Link:** <https://www.datakitchen.io/>

One of the most popular DataOps tools, DataKitchen is best for automating and coordinating people, environments, and tools in data analytics of the entire organization. DataKitchen handles it all – from testing to orchestration, to development, and deployment. Using this platform, your organization can achieve virtually zero errors and deploy new features faster than your business. DataKitchen lets organizations spin up repetitive work environments in a matter of minutes so teams can experiment without breaking production cycles. The Quality pipeline of DataKitchen is based on three core sections; data, production, and value. It is essential to understand that with this tool, you can access pipeline with Python Code, transform it via SQL coding, design model in R, visualize in Workbook, and gain reports in form of Tableau.

### Lentiq

**Website Link:** <https://lentiq.com/>

This data model deployment tool works in a service environment for smaller teams. With Lentiq, you can run data science and data analysis at the scale of your choice in the clouds so your team can ingest real-time data, process it, and share useful insights. With Lentiq, your team can train, build, and share models within the environment and innovate without restrictions. It is suggested to use Jupyter Notebooks for training models on Lentiq.

# Conclusion

Data Ops, or data operations, is an agile DevOps based methodology for design, implementation, and maintenance of data in a distributed architecture. The main goal of this approach is to offer efficient and accurate results on big data – received by organizations on a daily basis – and extract useful analytics.

All in all, DataOps are an entrance to the world of smarter products. Now organizations can utilize fully managed platforms to created autonomous data pipelines which not only fuel analytics but also ML applications. It is essential for companies to leverage DataOps platforms so that their teams can adopt and collaborate conveniently while working with enormous datasets on routine basis.